Supporting Information

Complementarity Between a Docking and High-Throughput Screen in Discovering New Cruzain Inhibitors

Rafaela S. Ferreira[1,2,3], Anton Simeonov[4]*, Ajit Jadhav[4],Oliv Eidam[2], Bryan T. Mott[4], James H. McKerrow[3], David J. Maloney[4] John J. Irwin[2] & Brian K. Shoichet[2]*

[1]Graduate Program in Chemistry and Chemical Biology, [2]Department of Pharmaceutical Chemistry and [3]Sandler Center for Basic Research in Parasitic Diseases, University of California San Francisco, CA, USA, 94158; [4]NIH Chemical Genomics Center, Bethesda, MD 20892-3370, USA

* corresponding authors

Table of Contents:

Supplementary Figure 7 Superposition of conformation of **27** in complex with cruzain and closest conformation found by docking (rmsd = 1.4 Å).

Supplementary Table S1 Follow up of qHTS Hits ranked among top 1% of the database by DOCK

Supplementary Table S2 DOCK ranking and experimental follow up of compound **4** analogues (Cluster 44)

Supplementary Table S3 Follow up of clusters selective for cruzain

Supplementary Table S4 Potential qHTS false negatives prioritized for testing by docking

Supplementary Table S5 *In vitro* activity of compound **11** and representative analogues

Supplementary Table S6 Comparison of DOCK scores for crystallographic and DOCK poses of compound **27**

Synthesis of analogues of compound 11

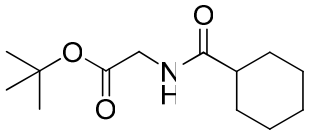
General Methods:

Unless otherwise stated, all reactions were carried out under an atmosphere of dry argon or nitrogen in dried glassware. Indicated reaction temperatures refer to those of the reaction bath, while room temperature (rt) is noted as 25 °C. All solvents were of anhydrous quality purchased from Aldrich Chemical Co. and used as received. Commercially available starting materials and reagents were purchased from Aldrich and were used as received.

Analytical thin layer chromatography (TLC) was performed with Sigma Aldrich TLC plates (5 x 20 cm, 60 Å, 250 μm). Visualization was accomplished by irradiation under a 254 nm UV lamp. Chromatography on silica gel was performed using forced flow (liquid) of the indicated solvent system on Biotage KP-Sil pre-packed cartridges and using the Biotage SP-1 automated chromatography system. ^{1}H- and ^{13}C NMR spectra were recorded on a Varian Inova 400 MHz spectrometer. Chemical shifts are reported in ppm with the solvent resonance as the internal standard ($CDCl_3$ 7.26 ppm, 77.00 ppm, DMSO-d_6 2.49 ppm, 39.51 ppm for ^{1}H, ^{13}C respectively). Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, br = broad, m = multiplet), coupling constants, and number of protons. Low resolution mass spectra (electrospray ionization) were acquired on an Agilent Technologies 6130 quadrupole spectrometer coupled to the HPLC system. High resolution mass spectral data was collected in-house using and Agilent 6210 time-of-flight mass spectrometer, also coupled to an Agilent Technologies 1200 series HPLC system. If needed, products were purified via a Waters semi-preparative HPLC equipped with a Phenomenex Luna® C18 reverse phase (5 micron, 30 x 75 mm) column having a flow rate of 45 mL/min. The mobile phase was a mixture of acetonitrile and H_2O each containing 0.1% trifluoroacetic acid. Samples were analyzed for purity on an Agilent 1200 series LC/MS equipped with a Luna® C18 reverse phase (3 micron, 3 x 75 mm) column having a flow rate of 0.8-1.0 mL/min over a 3-minute gradient and

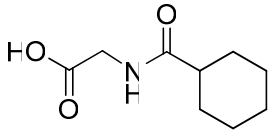
a 4.5 minute run time. The mobile phase was a mixture of acetonitrile (0.025% TFA) and H_2O (0.05% TFA), and a temperature was maintained at 50 ^oC.. Purity of final compounds was determined to be >95%, using a 3 µL injection with quantitation by AUC at 220 and 254 nM (Agilent Diode Array Detector).

Preparation of compound **31**:



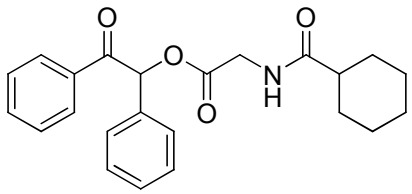
***tert*-butyl 2-(cyclohexanecarboxamido)acetate:**

To a solution of tert-butyl 2-aminoacetate (0.6 g, 4.57 mmol) in CH_2Cl_2 (25 mL) was added triethylamine (0.64 mL, 4.57 mmol) and cyclohexanecarbonyl chloride (0.67 g, 4.57 mmol). The reaction mixture was stirred for 1 h at rt, at which time the solution was washed with water. The organic layer was extracted, dried on $MgSO_4$, filtered, and concentrated in vacuo. The resulting residue was purified directly on silica column. Gradient elution with ethyl acetate (10→60%) in hexanes provided the title compound as a colorless solid: yield (1.0 g, 4.14 mmol, 91 %).



2-(cyclohexanecarboxamido)acetic acid:

To a solution of *tert*-butyl 2-(cyclohexanecarboxamido)acetate (1.0 g, 4.14 mmol) in toluene (100 mL) was added SiO_2 (20 grams). The reaction was heated to reflux overnight, then cooled and filtered. The SiO_2 was washed with 10% MeOH/CH_2Cl_2 several times. The filtrate was concentrated to yield the title compound as a colorless solid. No further purification was needed: yield (0.69 g, 3.73 mmol, 90%).

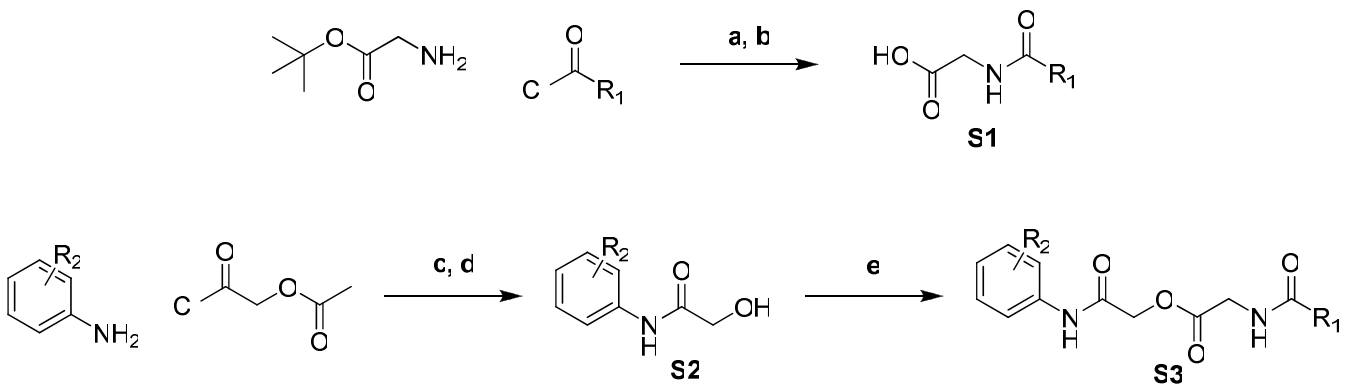


31

2-oxo-1,2-diphenylethyl 2-(cyclohexanecarboxamido)acetate:

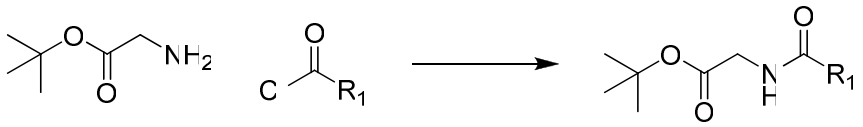
To a solution of commercially available 2-hydroxy-1,2-diphenylethanone (0.15 g, 0.71 mmol) and 2-(cyclohexanecarboxamido)acetic acid (0.13 g, 0.71 mmol) in THF (7 mL) was added triphenylphosphine (0.19 g, 0.71 mmol), followed by dropwise addition of diisopropyl azodicarboxylate (0.14 mL, 0.71 mmol). The reaction mixture was stirred at rt for 2 h. Upon completion, the solvent was removed under reduced pressure and the residue was purified directly on silica column. Gradient elution with ethyl acetate (1→35%) in hexanes provided the title compound as a colorless solid: yield (0.21 g, 0.55 mmol, 78 %). LC-MS: rt (min) = 3.68; ^{1}H NMR (DMSO-d_6) δ 1.06-1.36 (m, 6H), 1.52-1.73 (m, 4), 2.07-2.21 (m, 1H), 3.84-4.06 (m, 2H), 7.15 (s, 1H), 7.32-7.43 (m, 3H) 7.46-7.57 (m, 4H) 7.62 (t, *J* = 7.4 Hz, 1H) 8.05 (d, *J* = 7.2 Hz, 2H) 8.20 (t, *J* = 6.0 Hz, 1H); ^{13}C NMR (DMSO-d_6) δ 25.39, 25.55, 29.57, 43.91, 43.99, 44.06, 77.00, 78.19, 128.53, 128.72, 129.12, 129.77, 130.06, 133.73, 133.95, 134.26, 134.30, 134.82, 135.00, 170.02, 176.12, 176.29, and 193.92; HRMS (*m/z*): $[M]^+$ calcd. for $C_{23}H_{25}NO_4$, 379.1784; found, 379.1788.

General Scheme for compounds **32-35**:

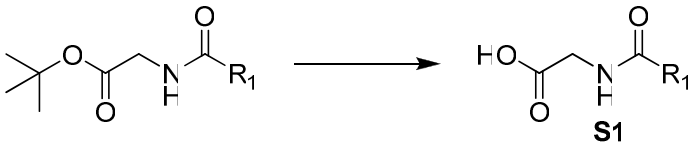


a) Et_3N, CH_2Cl_2, rt, 1 h; b) SiO_2, toluene, reflux, 16 h; c) Et_3N, CH_2Cl_2, rt, 1 h; d) K_2CO_3, MeOH 1 h; e) **1**, PPh_3, DIAD, THF, rt, 16 h.

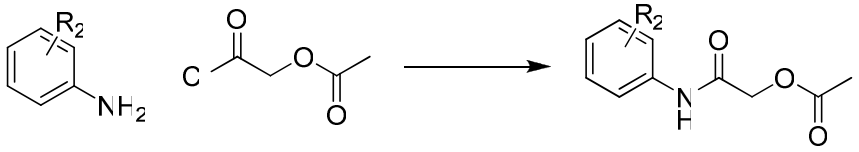
General Procedures:



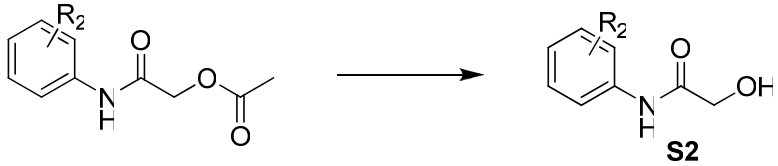
To a stirring solution of *t*-butyl glycine (1.0 eq) in CH_2Cl_2 were added the acid chloride (1.1 eq) and triethylamine (1.1 eq). The reaction mixture was stirred at rt for 1 hr, at which time the mixture was diluted further w/ CH_2Cl_2 and washed with sat. ammonium chloride solution and brine. The organic layer was extracted, dried on $MgSO_4$, filtered, and the solvent was removed under reduced pressure. The residue was purified directly on silica gel. Gradient elution (20-40% EtOAc in hexanes) afforded the desired product (*t*-butyl amidoacetates): yield (95-99%).



A solution of the *t*-butyl amidoacetate (1 mmol) and SiO_2 (6 g) in toluene was refluxed for 16 h. The reaction mixture was filtered and the silica gel was washed several times with 10% MeOH-CH_2Cl_2 and the solvent was removed under reduced pressure to afford the amidoacetic acids, **S1** as a colorless or pale solids: yield (90-95%). See below for specific example.

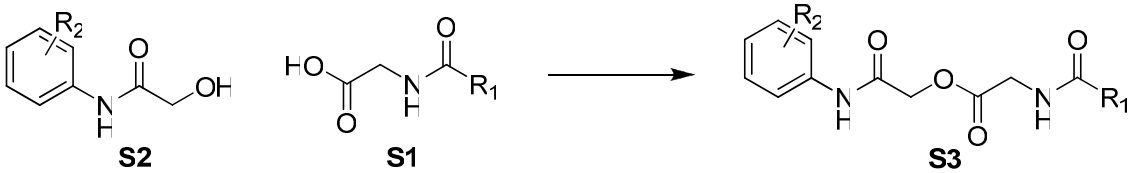


General procedure C – To a solution of the requisite aniline (1.0 eq) in CH_2Cl_2 were added triethylamine (1.1 eq) and 2-chloro-2-oxoethyl acetate (1.1 eq). The reaction mixture was stirred at rt for 1 h, at which time the mixture was diluted further w/ CH_2Cl_2 and washed with sat. ammonium chloride solution and brine. The organic layer was extracted, dried on $MgSO_4$, filtered, and the solvent was removed under reduced pressure. The residue was purified directly on silica gel. Gradient elution (20-40% EtOAc in hexanes) afforded the desired products (2-oxo-2-(arylamino)ethyl acetates) as a colorless or pale solid: yield (95-99%).



S2

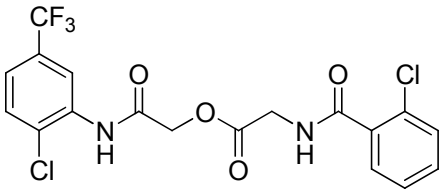
General procedure A – To a stirring solution of 2-oxo-2-(arylamino)ethyl acetate (1.0 eq) in a methanol with potassium carbonate (1.0 eq). The reaction mixture was stirred at rt for 1 h, at which time it was diluted w/ EtOAc and filtered through Celite. The organic layer was washed with sat. ammonium chloride and brine, extracted, dried on $MgSO_4$, filtered, and the solvent was removed under reduced pressure. The residue was purified directly on silica gel. Gradient elution (40-60% EtOAc in hexanes) afforded the desired product (2-hydroxy-*N*-arylacetamides, **S2**) as a colorless or pale solid or oil: yield (90-95%).



S2 **S1** **S3**

General procedure e – To a stirring solution of 2-hydroxy-*N*-arylacetamides **S2** (1.0 eq), amidoacetic acids **S1** (1.0 eq), and triphenylphosphine (1.1 eq) in THF was added diisopropyl

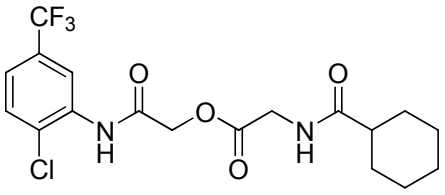
azodicarboxylate (DIAD, 1.1 eq) dropwise at rt. The reaction mixture was stirred for 16 h, at which time the solvent was removed under reduced pressure. The residue was purified directly on silica gel. Gradient elution (25-60% EtOAc in hexanes) afforded the desired products (benzamidoacetates or carboxamidoacetates, **S3**) as colorless or pale solids: yield (80-95%).



32

2-(2-chloro-5-(trifluoromethyl)phenylamino)-2-oxoethyl 2-(2-chlorobenzamido)acetate

LC-MS: rt (min) = 3.55; ^{1}H NMR ($CDCl_3$) δ 4.43 (d, 2H, *J* = 5.6 Hz), 4.87 (s, 2H), 6.98 (m, 1H), 7.33-7.44 (m, 4H), 7.52 (d, 1H, *J* = 8.4 Hz), 7.71 (d, 1H, *J* = 7.6 Hz), 8.56 (brs, 1H) and 8.65 (s, 1H); ^{13}C NMR ($CDCl_3$) δ 41.89, 63.67, 119.02 (q, F-splitting), 122.07 (q, F-splitting), 127.02, 127.67, 129.67, 130.38, 130.45, 130.55, 130.88, 131.99, 133.41, 134.06, 164.69, 166.73 and 168.17; HRMS (*m/z*): $[M]^+$ calcd. for $C_{18}H_{13}Cl_2F_3N_2O_4$, 448.0204; found, 448.0200.

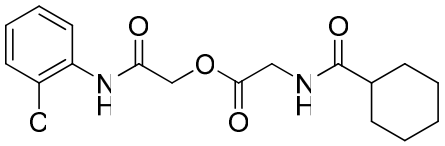


33

2-(2-chloro-5-(trifluoromethyl)phenylamino)-2-oxoethyl 2-(cyclohexanecarboxamido)acetate

LC-MS: rt (min) = 3.61; ^{1}H NMR (DMSO-d_6) δ 1.08-1.32 (m, 5H), 1.55-1.67 (m, 5H), 2.10-2.17 (m, 1H), 3.28 (s, 3H), 3.64 (s, 1H), 3.90 (d, 2H, *J* = 6.0 Hz), 7.53 (dd, 1H, *J* = 8.4 and 1.6 Hz), 7.75 (d, 1H, *J* = 8.4 Hz), 8.12 (d, 1H, *J* = 2.0 Hz), 8.15 (m, 1H) and 9.90 (s, 1H); ^{13}C NMR

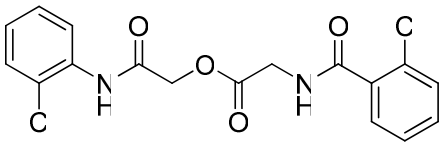
(DMSO-d_6) δ 25.14, 25.41, 29.03, 43.51, 51.91, 60.74, 62.59, 121.73 (q, F-splitting), 122.17, 122.85 (q, F-splitting), 124.89, 127.89, 128.22, 130.15, 130.87, 135.10, 166.34, 169.69, 175.64 and 175.80; HRMS (*m/z*): $[M]^+$ calcd. for $C_{18}H_{20}ClF_3N_2O_4$, 420.1064; found, 420.1068.



34

2-(2-chlorophenylamino)-2-oxoethyl 2-(cyclohexanecarboxamido)acetate

LC-MS: rt (min) = 3.38; ^{1}H NMR (DMSO-d_6) δ 1.1-1.35 (m, 5H), 1.58-1.70 (m, 5H), 2.15 (tt, 1H, *J* = 11.2 and 3.2 Hz), 3.93 (d, 2H, *J* = 6.0 Hz), 4.76 (s, 2H), 7.21 (td, 1H, *J* = 7.6 and 1.6 Hz), 7.33 (td, 1H, *J* = 8.0 and 1.6 Hz), 7.50 (dd, 1H, *J* = 8.0 and 1.6 Hz), 7.68 (dd, 1H, *J* = 8.0 and 1.6 Hz) and 8.20 (t, 1H, *J* = 6.0 Hz); ^{13}C NMR (DMSO-d_6) δ 25.14, 25.42, 29.04, 43.53, 126.21, 126.74, 127.49, 129.53, 134.06, 165.74, 169.64 and 175.81; HRMS (*m/z*): $[M]^+$ calcd. for $C_{17}H_{21}ClN_2O_4$, 352.1190; found, 352.1187.



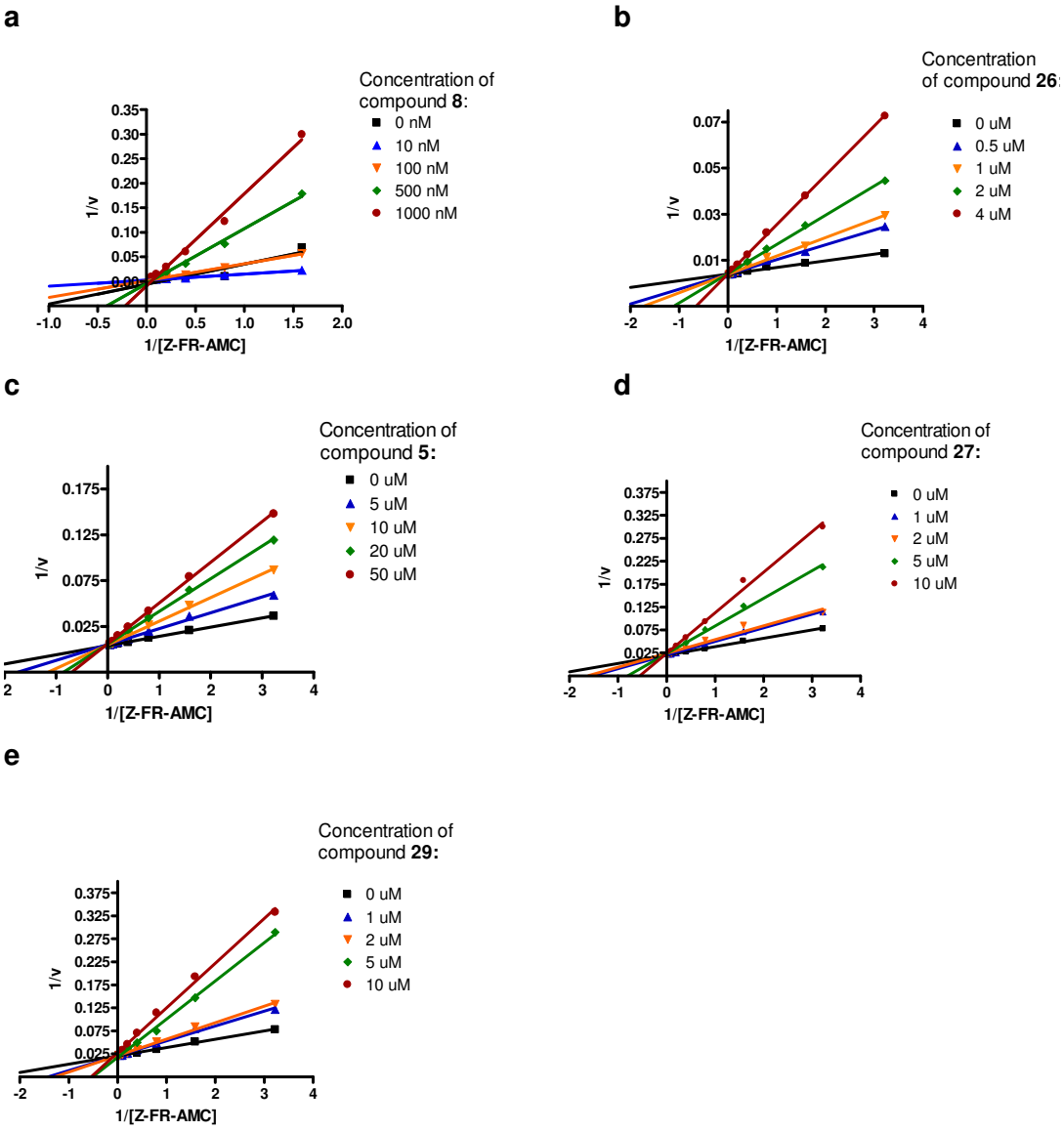
35

2-(2-chlorophenylamino)-2-oxoethyl 2-(2-chlorobenzamido)acetate

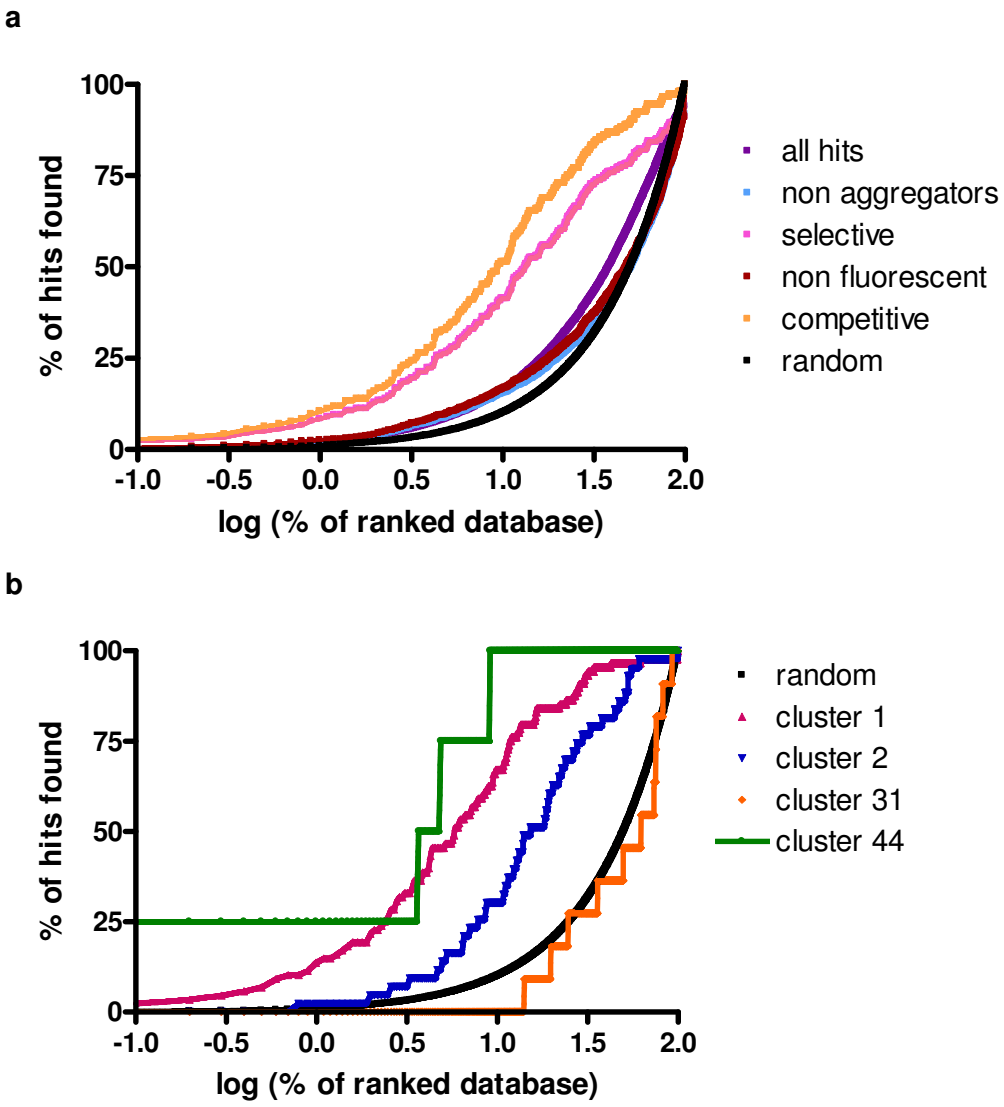
LC-MS: rt (min) = 3.37; ^{1}H NMR (DMSO-d_6) δ 4.15 (d, 2H, *J* = 6.0 Hz), 4.83 (s, 2H), 7.22 (td, 1H, *J* = 7.6 and 1.6 Hz), 7.33 (td, 1H, *J* = 8.0 and 1.2 Hz), 7.38-7.52 (m, 5H), 7.69 (dd, 1H, *J* = 8.0 and 1.6 Hz), 8.95 (t, 1H, *J* = 6.0 Hz) and 9.70 (s, 1H); ^{13}C NMR (DMSO-d_6) δ 40.79, 62.72, 126.24, 126.76, 127.07, 127.49, 129.02, 129.54, 129.73, 129.99, 131.09, 134.06, 135.97,

165.69, 166.85 and 169.05; HRMS (*m/z*): $[M]^+$ calcd. for $C_{17}H_{14}Cl_2N_2O_4$, 380.0331; found, 380.0331.

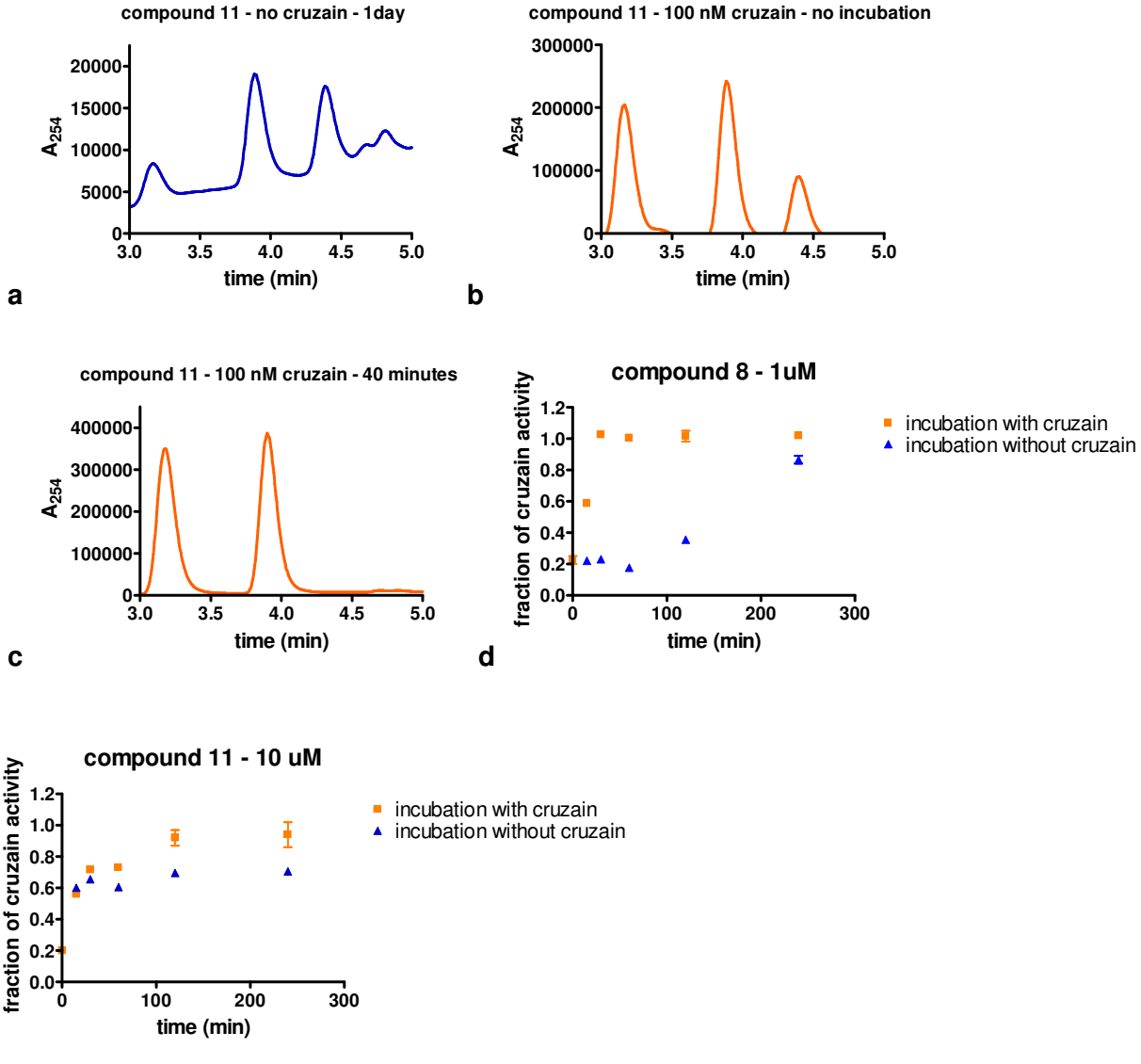
Supplementary Figure 1 Lineweaver-Burk plots for representative compounds for five classes of cruzain competitive inhibitors. (a) **8**, cluster 1, apparent Ki =65 nM (b) **26**, Ki = 0.8 μM, (c) **5**, Ki = 6 μM, (d) **27**, cluster 2, Ki = 2 μM, (e) **29**, cluster 31, Ki = 2 μM.



Supplementary Figure 2 Enrichment curves. (a) Improvement of enrichment at each stage of mechanistic follow up. Curves for all hits (purple), detergent insensitive compounds (putative non aggregators) (cyan), non fluorescent compounds (red), compounds selective for cruzain (pink) and competitive inhibitors (orange) are shown. (b) Enrichments for each cluster of competitive inhibitors. Clusters 1, 2, 31 and 44 are shown in magenta, blue, orange and green respectively. Enrichment expected by random ranking of compounds shown in black.

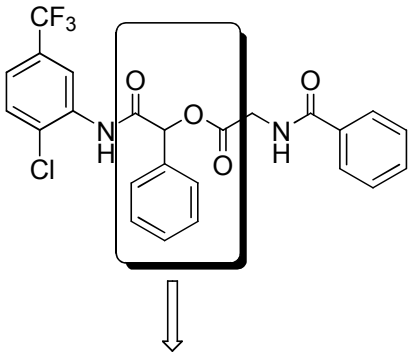


Supplementary Figure 3 Cruzain catalyzed degradation of cluster 1 compounds. UV trace for absorbance at 254 nm for solutions of **11** (a) solution after 1 day in the absence of cruzain, (b) fresh solution in presence of 100 nM cruzain, (c) after 40 minutes in the presence of 100 nM cruzain. The peak eluted at 4.4 minutes referes to compound **11**, whereas the one eluted at 3.2 min corresponds to a product of **11** cleavage. Time-dependence of cruzain inhibition by compounds (d) **8** and (e) **11**.



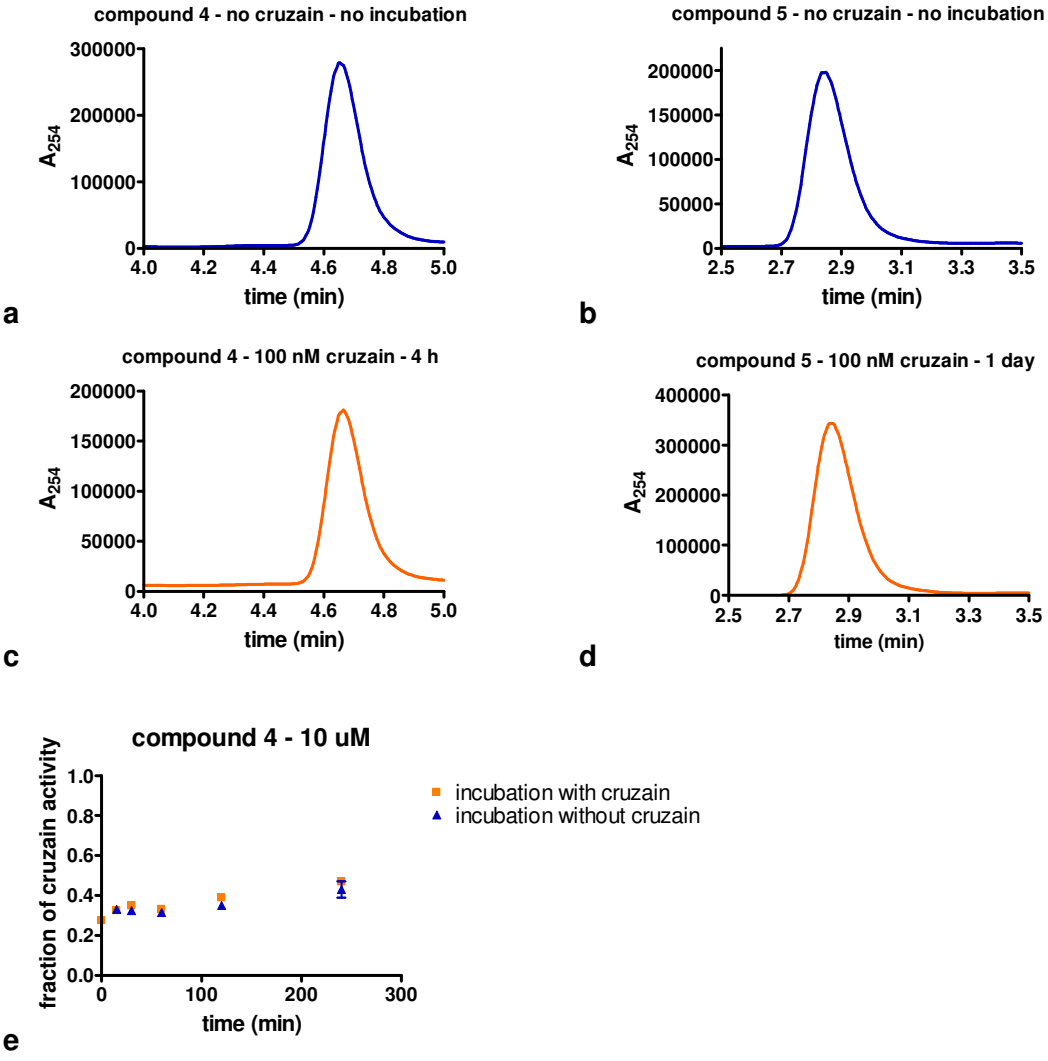
Supplementary Figure 4 Replacement of ester functionality in **11** yields inactive compounds.

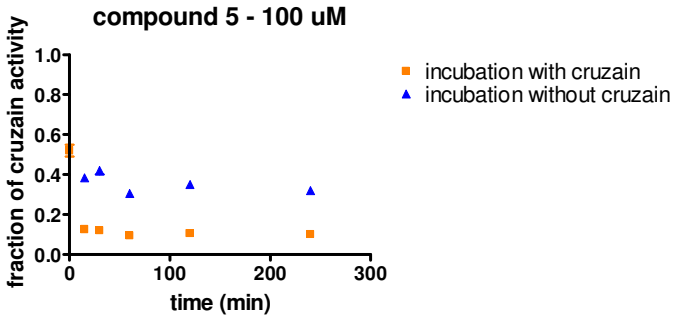
NO ACTIVITY WAS OBSERVED



replaced ester with amide, oxadizaole, thiadiazole triazole.

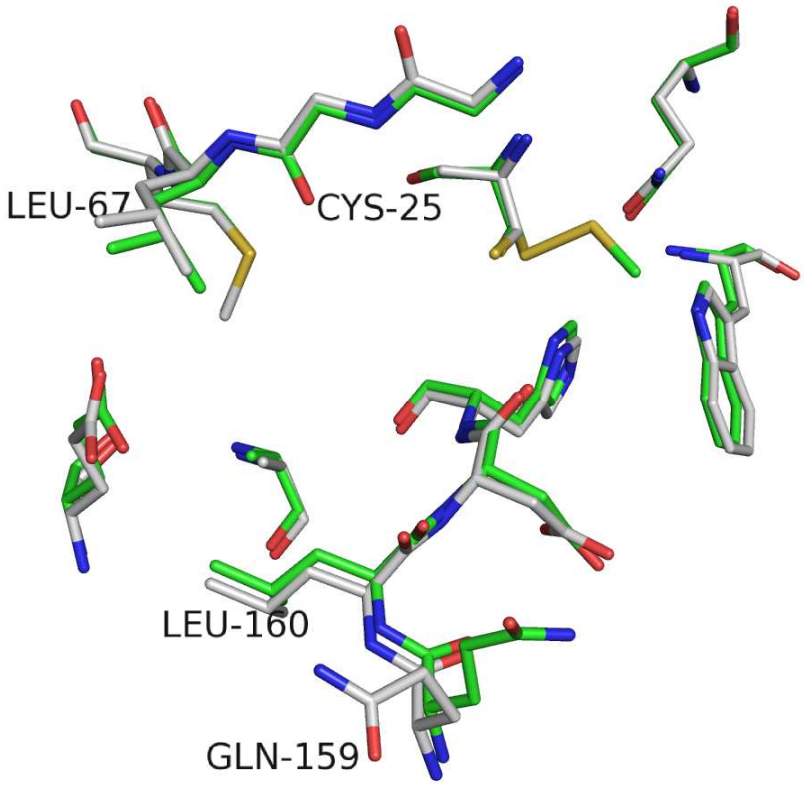
Supplementary Figure 5 Chemical stability and time-dependence cruzain inhibition data for compounds **4** and **5**. UV trace for absorbance at 254 nm for fresh compound solutions of (a) **4** and (b) **5** in absence of cruzain**;** and for solutions incubated with 100 nM cruzain: (c) **4** after 4 h incubation and , (d) **5** after 1 day incubation. Time-course of cruzain inhibition by compounds (e) **4** and (f) **5**, indicating no time-dependence over 240 min incubation.



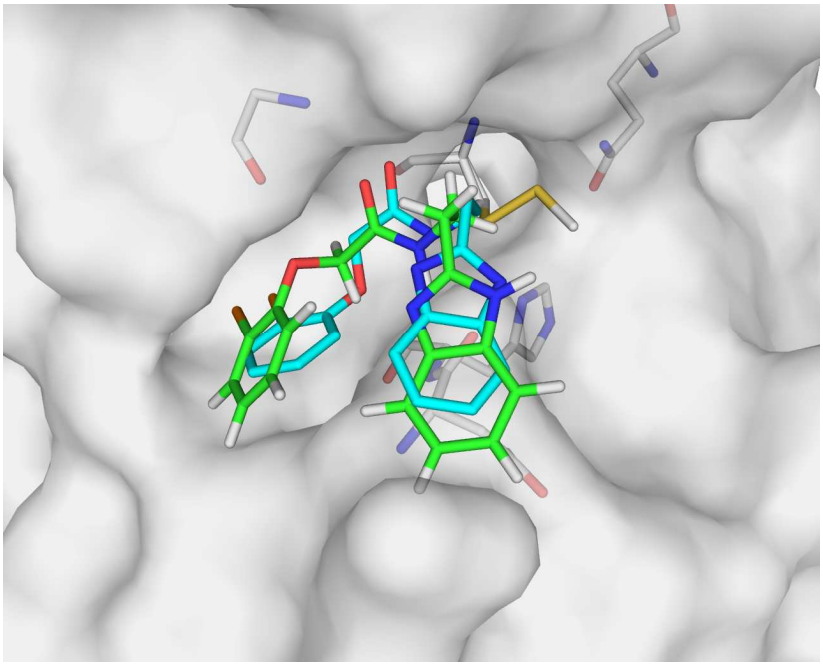

compound 5 - 100 uM
fraction of cruzain activity
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
time (min)
incubation with cruzain
incubation without cruzain

f

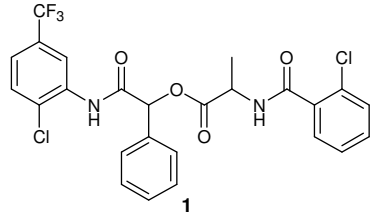
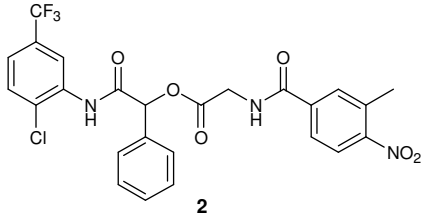
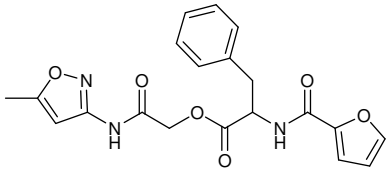
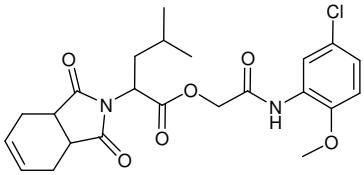
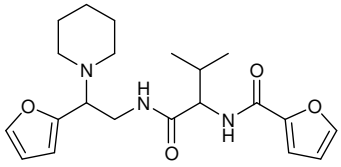
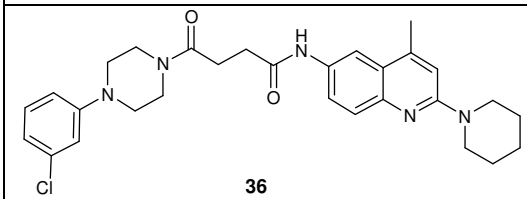
Supplementary Figure 6 Comparison between cruzain structure used for docking and crystal structure of cruzain/**27** complex. Residues within 5 Å of compound **27** are shown in sticks. Conformations of active site residues are similar in both structures, except for Gln159 and double conformation of Cys25. Carbon atoms colored green in crystal structure and gray in structure used for docking. Oxygen, nitrogen and sulfur colored red, blue and yellow respectively. Figures prepared with Pymol.[24]

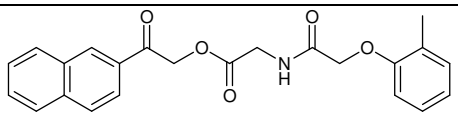
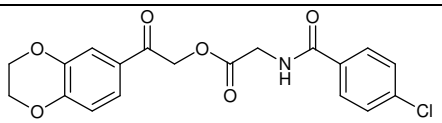
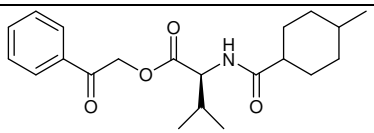
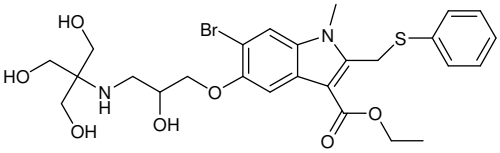
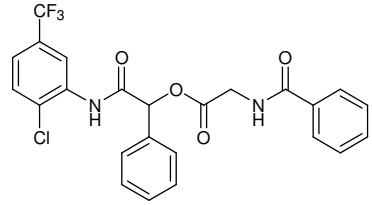
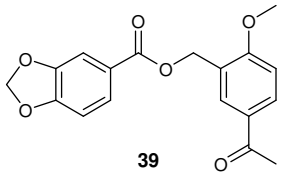
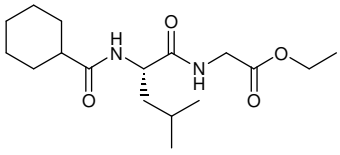
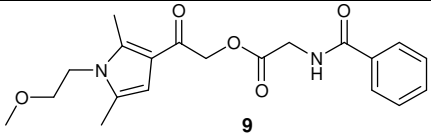


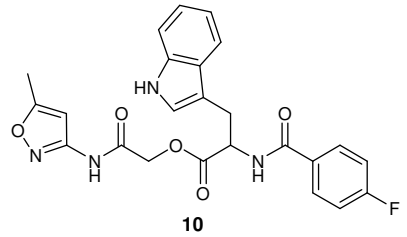
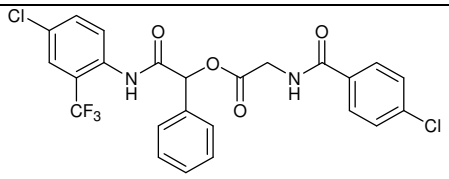
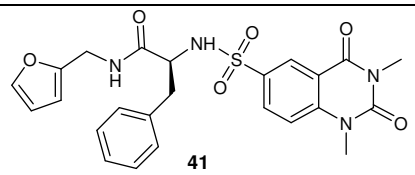
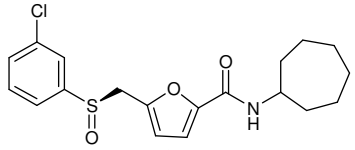
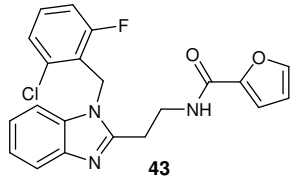
Supplementary Figure 7 Superposition of conformation of **27** in complex with cruzain and closest conformation found by docking (rmsd = 1.4 Å). Carbon atoms colored cyan in crystallographic complex and green in docked conformation. Oxygen, nitrogen, sulfur and bromine colored red, blue, yellow and orange respectively. Figures prepared with Pymol.[24]



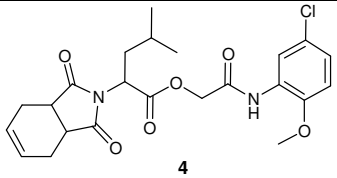
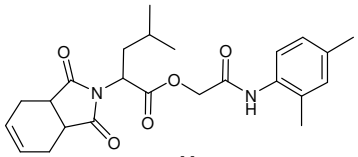
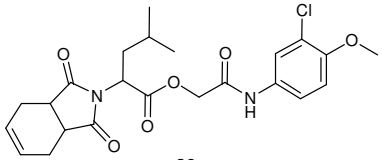
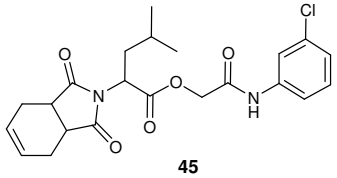
Supplementary Table S1 - Follow up of qHTS Hits ranked among top 1% of the database by DOCK

Structure	DOCK rank	IC50 (μM)	AmpC inhibition?	Detergent-sensitivity?	Time-dependence?
 **1**	6	11	No	No	No
 **2**	20	0.4	No	No	No
 **3**	97	38	No	No	No
 **4**	153	1	No	No	No
 **5**	173	7	No	No	No
 **36**	311	25	Yes	No	No

 **6**	550	0.7	No	No	No
 **7**	555	0.5	No	No	No
 **37**	647	166	Yes	No	No
 **38**	734	-	No	Yes	No
 **8**	789	0.3	No	No	No
 **39**	951	-	yes	Yes	yes
 **40**	1151	65	no	No	No, but low inhibition
 **9**	1182	18	No	No	No

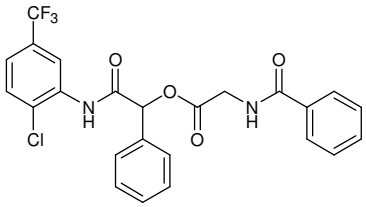
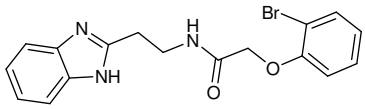
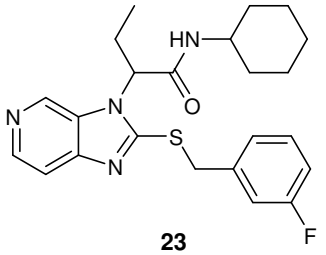
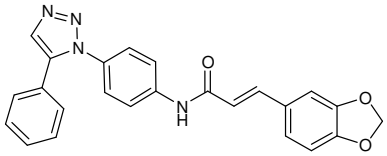
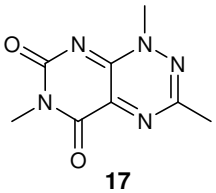
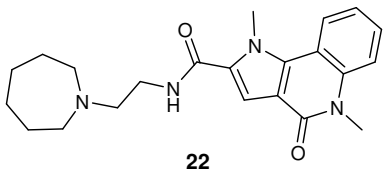
 **10**	1378	3	No	No	No
 **11**	1485	0.7	No	No	No
 **41**	1623	-	No	Yes	No
 **42**	1825	-	No	Yes	Yes
 **43**	1833	-	No	Yes	No

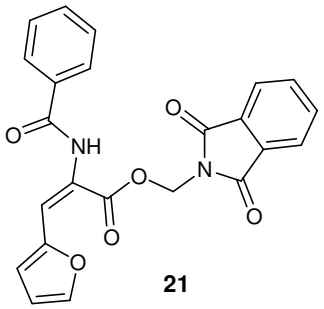
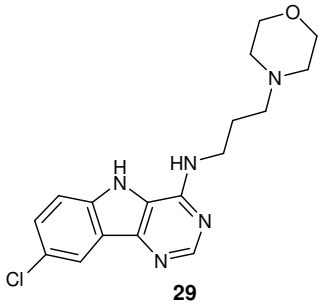
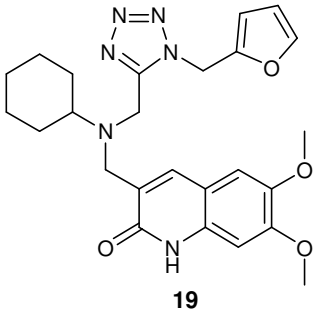
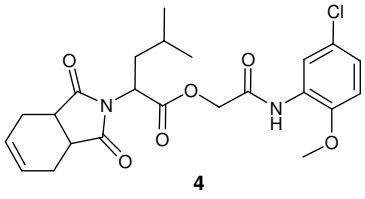
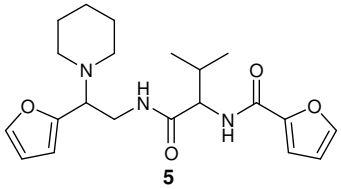
Supplementary Table S2 – DOCK ranking and experimental follow up of compound **4** analogues (Cluster 44)

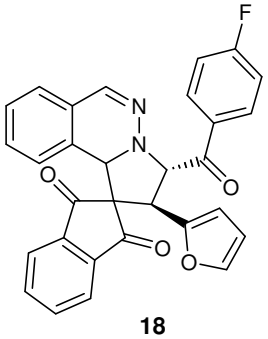
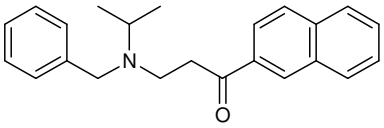
Structure	DOCK rank	IC_{50} (μM)	AmpC inhibition?	Detergent-sensitivity?	Time-dependence?	Ki (μM)
 **4**	153	1	No	No	No	1.6
 **44**	8593	25	ND	No	No	ND
 **26**	11337	0.9	No	No	No	0.8
 **45**	21010	5	ND	No	No	ND

ND = not determined

Supplementary Table S3 Follow up of clusters selective for cruzain

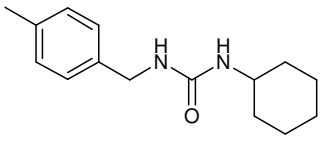
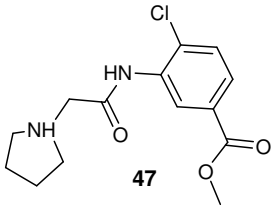
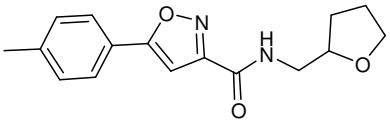
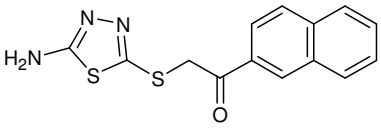
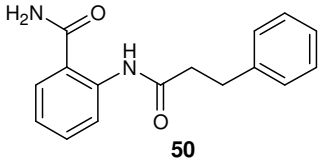
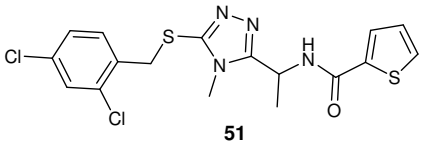
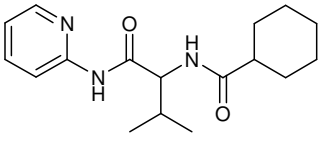
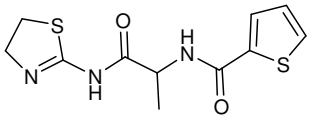
Cluster/ compounds per cluster	Compound tested experimentally	qHTS IC_{50} (μM)	Detergent sensitive?	Time-dependent ?	% β-lactamase inhibition
1/88	 **8**	7	No	no	0 (10 μM)[a]
2/43	 **27**	0.4	No	no	0 (100 μM)[a]
12/8	 **23**	6	yes	inconclusive	NA
21/2	 **20**		No	yes	NA
27/5	 **17**	2	No	yes	NA
28/7	 **22**	25	Yes	no	NA

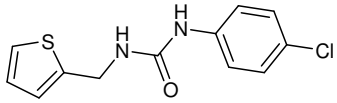
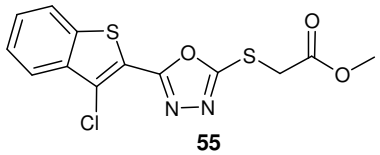
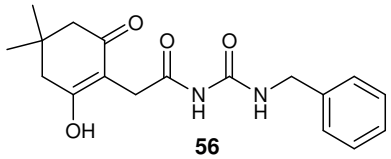
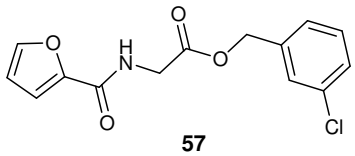
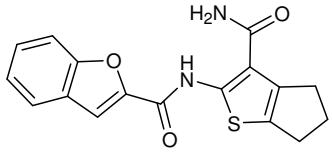
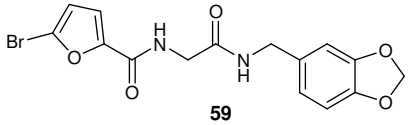
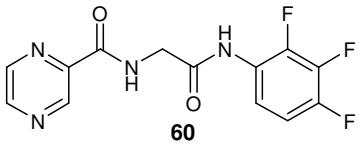
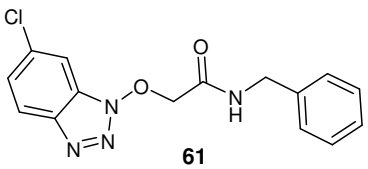
30/5	 **21**	1	No	yes	NA
31/10	 **29**	13	No	no	0 (100 μM)[a]
37/3	 **19**	14	No	yes	NA
44/4	 **4**	13	No	no	5 (10 μM)[a]
Singleton	 **5**	14	No	no	6 (10 μM)[a]

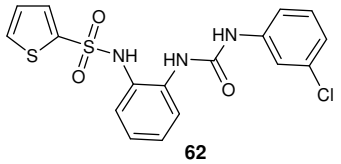
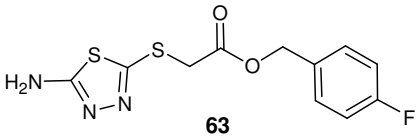
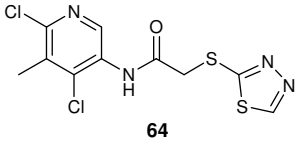
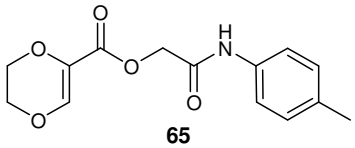
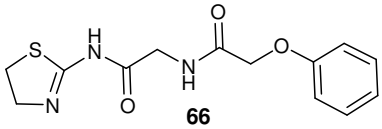
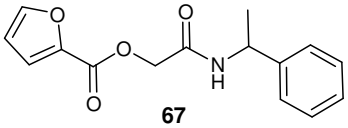
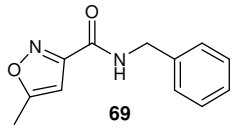
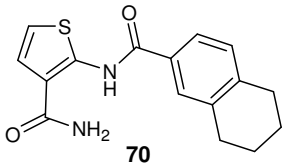
Singleton	 **18**	8	No	yes	NA
Singleton	 **24**	32	No	no	NA

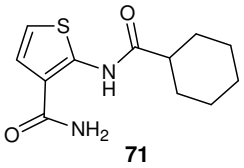
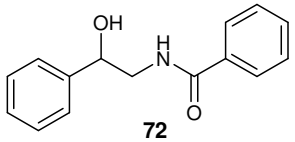
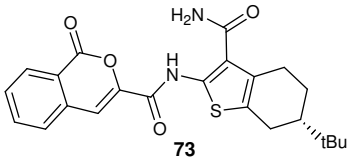
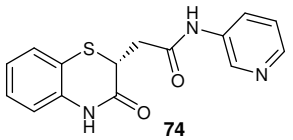
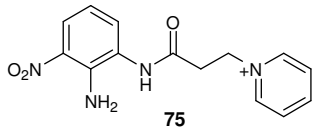
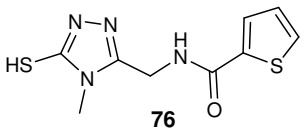
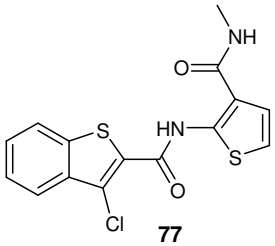
[a] compound concentration in assay.

Supplementary Table S4 – Potential qHTS false negatives prioritized for testing by docking

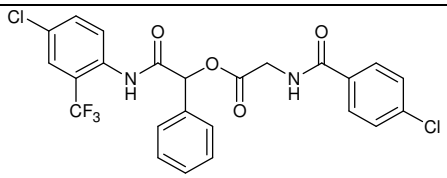
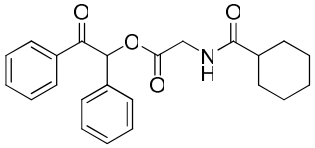
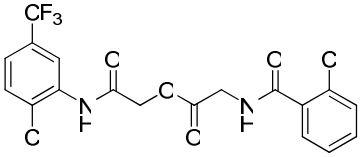
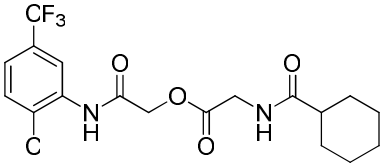
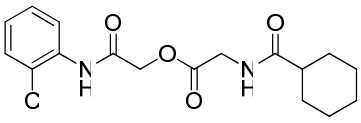
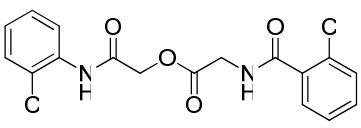
Structure	% cruzain inhibition at 200 μM
 **46**	0
 **47**	0
 **48**	0
 **49**	23
 **50**	0
 **51**	0
 **52**	0
 **53**	6

Structure	% cruzain inhibition at 200 μM
 **54**	1
 **55**	**49**
 **56**	8
 **57**	25
 **58**	0
 **59**	0
 **60**	7
 **61**	19

Structure	% cruzain inhibition at 200 μM
 **62**	**74**
 **63**	32
 **64**	**62**
 **65**	3
 **66**	25
 **67**	2
 **69**	**42**
 **70**	High fluorescence

Structure	% cruzain inhibition at 200 μM
 **71**	8
 **72**	8
 **73**	**68**
 **74**	**40**
 **75**	**50**
 **76**	8
 **77**	60

Supplementary Table 5 *In vitro* activity of compound **11** and representative analogues

Compound	IC_{50} (nM) against Cruzain
 **11**	260
 **31**	30
 **32**	220
 **33**	670
 **34**	520
 **35**	250

Supplementary Table 6 Comparison of DOCK scores for crystallographic and DOCK poses of compound **27**

Cruzain structure	**Compound 27 pose**	**Scores**			
		electrostatics	**van der Waals**	**ligand desolvation**	**Total score**
Same used for virtual screening	Best ranked by docking	- 33.0	- 22.2	19.7	- 35.5
	Crystallographic	- 17.0	- 8.4	16.9	- 8.5
From structure complex with **27**	Best ranked by docking	- 12.06	- 25.9	11.2	- 26.7
	Crystallographic	- 11.7	- 22.5	16.6	- 17.8